Nonlocal setting and outcome information for violation of Bell's inequality

Marcin Pawłowski,[1,2] Johannes Kofler,[1,3] Tomasz Paterek,[1,4] Michael Seevinck,[5] and Časlav Brukner[1,3]

[1]*Institute for Quantum Optics and Quantum Information, Austrian Academy of Sciences, Boltzmanngasse 3, A-1090 Vienna, Austria*
[2]*Institute of Theoretical Physics and Astrophysics, University of Gdańsk, ul. Wita Stwosza 57, PL-80-952 Gdańsk, Poland*
[3]*Faculty of Physics, University of Vienna, Boltzmanngasse 5, A-1090 Vienna, Austria*
[4]*Centre for Quantum Technologies and Department of Physics, National University of Singapore, 3 Science Drive 2, 117542 Singapore, Singapore*
[5]*Institute for Theoretical Physics, Utrecht University, P.O.Box 80.195, 3508 TD Utrecht, the Netherlands.*

Bell's theorem is a no-go theorem stating that quantum mechanics cannot be reproduced by a physical theory based on realism, freedom to choose experimental settings and two locality conditions: setting (SI) and outcome (OI) independence. We provide a novel analysis of what it takes to violate Bell's inequality within the framework in which both realism and freedom of choice are assumed, by showing that it is impossible to model a violation without having information in one laboratory about *both* the setting and the outcome at the distant one. While it is possible that outcome information can be revealed from shared hidden variables, the assumed experimenter's freedom to choose his settings forces that setting information *must be* non-locally transferred, even when SI is obeyed. The sufficient amount of transmitted information about the setting to violate the CHSH inequality up to its quantum mechanical maximum is 0.736 bits.

PACS numbers: 03.65.Ta, 03.65.Ud

Bell's inequalities are certain constraints on correlations between space-like separated measurements which are satisfied in any local-realistic theory [1]. The inequalities are violated by quantum predictions for some entangled states. The usual set of assumptions invoked in the derivation of Bell's inequalities are *realism*, the experimenter's *freedom* to choose the measurement settings ("freedom of choice"), and two *locality* conditions: setting independence and outcome independence [2–4]. Maintaining realism and freedom of choice thus seems to necessitate an exchange of information between distant measurement stations that defies locality so as to violate one (or both) of the independence conditions. What kind of information and which amount of it has to be transferred between the stations to model the violation of Bell's inequalities? Is it information about the distant outcome, or about the distant setting, or about both? These questions will be addressed in the present Letter thereby providing a novel analysis of what it takes to violate Bell's inequality.

In addition to fundamental reasons, answering the above questions is important in quantum information, such as e.g. in quantum communication complexity problems [5]. The question how much setting and/or outcome information needs to be exchanged in a Bell experiment for a given degree of violation is thus relevant for quantifying the classical resources required to simulate quantum efficiency in these problems.

Terhal *et al.* [6] have already shown that it is sufficient for a simulation of Bell experiments within non-local hidden-variable models (in the one-way communication model) to send the distant setting and let the hidden variable be the distant outcome. In such a case information about both the distant setting and the distant outcome are locally available. In this Letter we show that this is not just sufficient but also *necessary*: in order to obtain a violation of Bell's inequality it *must* be the case that *both* information about the measurement outcome and the measurement setting at one laboratory are available at the distant laboratory. This is to be contrasted with the well-known fact [3, 4] that either outcome independence or setting independence for (hidden-variable) conditional probabilities *can still be obeyed* in models giving a violation of the CHSH inequality, thereby indicating the novelty of our analysis of what it takes to violate the Bell inequality.

It is furthermore shown that, while it is possible in some models that information about the distant outcome can be read from the hidden-variables received from the source, the information about the setting *must* be non-locally transmitted, implicitly or explictly, in any model where the experimenters are free to choose their settings. We are able to trace this asymmetry between setting and outcome information to the freedom of the experimenters to choose their settings. We furthermore apply our analysis to the non-local hidden-variable models of Toner and Bacon [7], Leggett [8] and Bohm [9]. Finally, we show that the sufficient amount of transmitted information about the setting to violate the CHSH inequality up to its quantum mechanical maximum is 0.736 bits.

We begin with the usual formal definitions of the assumptions of Bell's theorem. In our notation a and b stand for the measurement settings chosen by the two distant experimenters, conventionally called Alice and Bob, respectively; A and B denote their respective measurement outcomes, and λ denotes a set of hidden variables.

(i) For stochastic (probabilistic) hidden-variable theories the assumption of *realism* dictates that the hidden variable λ specifies joint (non-negative, properly normalized) probabilities $P(A_{1,1}, A_{1,2}, A_{2,1}, ..; B_{1,1}, B_{1,2}, B_{2,1}, ..|\lambda)$, where e.g. the result $A_{1,2}$ of Alice for her setting choice 1 can depend on some non-local parameter 2, typically Bob's setting choice. The conditional probabilities $P(A, B|a, b, \lambda)$ that will be used in this Letter are then obtained as marginals of these joint ones.

(ii) *Setting Independence* (SI), often also called Parameter Independence [3], is the part of the locality condition

which prohibits the dependence of the probability to obtain the outcome in one laboratory on the choice of the setting at the other one: $P(A|a,b,\lambda) = P(A|a,\lambda)$, and analogous for $P(B|\cdot)$. Similarly, under *Outcome Independence* (OI), Alice's probability to obtain her outcome does not depend on Bob's outcome and vice versa: $P(A|a,b,B,\lambda) = P(A|a,b,\lambda)$, again analogous for $P(B|\cdot)$. The conjunction of these two conditions is equivalent to Bell's condition of Local Causality [2–4, 10]: $P(A,B|a,b,\lambda) = P(A|a,\lambda)P(B|b,\lambda)$. The latter condition allows to define the joint probabilities from (i) as $p(A_1,A_2|\lambda)p(B_1,B_2|\lambda)$, where e.g. A_1 is the result of Alice for her setting choice 1 [11].

(iii) The experimenter's *freedom of choice* to choose the measurement setting imposes that the selected measurement setting is statistically independent of the hidden variables sent by the source (even in a deterministic model). In terms of the (Shannon) mutual information this assumption is expressed as $I(\lambda : a) = I(\lambda : b) = 0$. As we will show, the assumption of freedom of choice is responsible for the *fundamental asymmetry* between settings and outcomes because it guarentees that the settings, contrary to the outcomes, are to be considered as independent variables.

Under these three assumptions the Clauser-Horne-Shimony-Holt (CHSH) inequality [12] must be obeyed:

$$\frac{1}{4}\sum_{a,b=0}^{1} P(A \oplus B = ab|a,b) \leq \frac{3}{4}, \qquad (1)$$

with $\oplus$ denoting addition modulo 2. We let Alice and Bob each choose with 50 % probability one of two settings, $a,b = 0,1$, and obtain measurement results, $A,B = 0,1$, respectively.

Assuming freedom of choice and realism, violations of the CHSH inequality imply that either OI or SI, or both, needs to be given up. In the framework of 'experimental metaphysics' [3] it is violation of the condition OI that is supposed to be responsible for the violation of the CHSH inequality, and it is argued that this is not an instance of action at a distance but only of some innocent 'passion at a distance': one passively comes to know the faraway outcome, but one cannot actively change it. However, upon closer scrutiny, both SI and OI are in fact conditions about probabilities for the *local* outcome only. Indeed, these conditions do not address all the possibilities, and in particular that information about *distant* settings and outcomes can be non-locally inferred. Here we will provide such an analysis.

We will show that within the framework of non-local realistic theories it is impossible to model a violation of the CHSH inequality without having information in one laboratory about *both* the setting and the outcome at the distant one. Thus, the increase in non-local information displayed by models that violate the CHSH inequality is necessarily about *both* the non-local settings and the outcomes, despite the fact that it is *not necessary* that the models are *both* setting dependent ($\neg$ SI) *and* outcome dependent ($\neg$ OI).

In order to proof our results we consider a local hidden-variable model augmented with information available to Alice about Bob's laboratory. While it is not necessary, it is instructive to think about this information as one-way classical communication from Bob to Alice. In every run of the experiment, Alice and Bob first choose their settings (a and b) and receive hidden variables λ which are independent of the choice of the settings. Then, Bob (or some process in his lab) generates the outcome B which in general depends on λ and b. Next, Bob generates the *message* $\mathcal{X}$ which depends on λ, b and B. Both the generation of B and of $\mathcal{X}$ are, in general, probabilistic processes. It is assumed that the exact mechanism how B and $\mathcal{X}$ are generated is known to Alice. Finally, $\mathcal{X}$ is transmitted to Alice who uses her optimal strategy, based on the knowledge of her setting a, the shared hidden variables λ, Bob's mechanisms and the message $\mathcal{X}$, to produce her outcome A in order to maximally violate the CHSH inequality.

From Alice's perspective, the CHSH inequality reads

$$\frac{1}{2}P(A = B|a = 0) + \frac{1}{2}P(A = B \oplus b|a = 1) \leq \frac{3}{4}, \qquad (2)$$

where, e.g. $P(A = B|a = k)$ is the probability that the outcome of Alice equals that of Bob, given she has chosen the kth setting. We shall show that the probabilities entering Ineq. (2) can be interpreted as a measure of the information Alice has about Bob's measurement setting and outcome. For this aim, we introduce the "guessed information" (GI):

$$J(\mathcal{X} \to \mathcal{Y}) := \sum_i P(\mathcal{X} = i)\max_j[P(\mathcal{Y} = j|\mathcal{X} = i)], \qquad (3)$$

where $\mathcal{X}$ takes values $i = 1,...,X$ and $\mathcal{Y}$ values $j = 1,...,Y$. The value of $J(\mathcal{X} \to \mathcal{Y})$ gives the average probability to correctly guess $\mathcal{Y}$ knowing the value of $\mathcal{X}$. Its maximum is 1 and then $\mathcal{Y}$ is fully specified by $\mathcal{X}$. The minimum of $J(\mathcal{X} \to \mathcal{Y})$ equals $\frac{1}{Y}$ and then $\mathcal{X}$ reveals no information about $\mathcal{Y}$. We note that GI reaches its minimum when the mutual information is $I(\mathcal{X} : \mathcal{Y}) = 0$, and it is maximal when $I(\mathcal{X} : \mathcal{Y}) = \log Y$. As an example, freedom of choice can be stated as $J(\lambda \to a,b) = \frac{1}{4}$, i.e., λ cannot reveal any information about the settings a and b. This implies the weaker condition $J(\lambda \to b) = \frac{1}{2}$.

Alice now uses an optimal maximization strategy so as to maximally violate the CHSH inequality. Consider the case in which Alice chooses $a = 0$. Her goal is to maximize the probability $P(A = B|a = 0)$ given the communicated value $\mathcal{X}$ and the received hidden variables λ. This maximized probability is just the average probability to correctly guess B given $\mathcal{X}$ and λ: $J(\lambda,\mathcal{X} \to B)$. Similarly, if her setting is $a = 1$, the maximal probability $P(A = B \oplus b|a = 1)$ equals $J(\lambda,\mathcal{X} \to B \oplus b)$. This allows to phrase the CHSH inequality in terms of the GI's

$$\frac{1}{2}J(\lambda,\mathcal{X} \to B) + \frac{1}{2}J(\lambda,\mathcal{X} \to B \oplus b) \leq \frac{3}{4}. \qquad (4)$$

We are now in the position to prove that a *necessary* condition for the violation of Bell's inequalities within non-local realism is that both information about the setting and about the outcome produced at one lab must be available at the distant lab. If there is no outcome information available, i.e.,

$J(\lambda, \mathcal{X} \to B) = \frac{1}{2}$, the left-hand side of Ineq. (4) cannot exceed $\frac{3}{4}$. To prove that setting information is also necessary, note that if one knows both B and $B \oplus b$, one also knows b. Thus, the average probability of correctly guessing b is greater or equal to the product of the average probabilities for the correct guess of B and $B \oplus b$:

$$J(\lambda, \mathcal{X} \to b) \geq J(\lambda, \mathcal{X} \to B) J(\lambda, \mathcal{X} \to B \oplus b). \qquad (5)$$

If $\mathcal{X}$ and λ carry no information about the setting, i.e., $J(\lambda, \mathcal{X} \to b) = \frac{1}{2}$, Ineq. (5) can be rewritten as $J(\lambda, \mathcal{X} \to B \oplus b) \leq \frac{1}{2} J^{-1}(\lambda, \mathcal{X} \to B)$, which implies

$$\frac{1}{2} J(\lambda, \mathcal{X} \to B) + \frac{1}{2} J(\lambda, \mathcal{X} \to B \oplus b) \leq \frac{1}{2} J(\lambda, \mathcal{X} \to B) + \frac{1}{4 J(\lambda, \mathcal{X} \to B)}. \qquad (6)$$

for the left-hand side of (4). This value is less or equal $\frac{3}{4}$ for the whole range of $J(\lambda, \mathcal{X} \to B) \in [\frac{1}{2}, 1]$. Thus, if there is no setting information, the violation of Ineq. (4), or Ineq. (1), is impossible.

Although both the information about the distant setting and about the distant outcome must be available at the local laboratory to have a violation, we show that, given freedom of choice, the information about the distant setting has to be transmitted non-locally, whereas it is possible that the information about the distant outcome can be obtained without any transmission from the shared hidden variables. This is shown by a further analysis of what information has to be transmitted via the message $\mathcal{X}$, over and above the information in the hidden variable λ. To this end, we introduce a measure of information, that we call "transmitted information" (TI), which is the difference of the averaged probability of correctly guessing the value of the variable $\mathcal{Y}$ when knowing $\mathcal{X}$ and λ, and the one when knowing only λ:

$$\Delta_\lambda(\mathcal{X} \to \mathcal{Y}) := J(\lambda, \mathcal{X} \to \mathcal{Y}) - J(\lambda \to \mathcal{Y}). \qquad (7)$$

$\Delta_\lambda(\mathcal{X} \to \mathcal{Y})$ takes values between 0 and $1 - \frac{1}{Y}$. Its lowest value means that transmission of $\mathcal{X}$ does not increase Alice's chances of guessing the correct value of $\mathcal{Y}$; $\mathcal{X}$ carries no *new* information about $\mathcal{Y}$ that is not already available to Alice through λ.

We have already established that either $J(\lambda, \mathcal{X} \to B) = \frac{1}{2}$ or $J(\lambda, \mathcal{X} \to b) = \frac{1}{2}$ implies no violation of the CHSH inequality. The asymmetry between the outcome and setting information originates from the freedom of choice assumption $J(\lambda \to b) = \frac{1}{2}$, which leads to

$$J(\lambda, \mathcal{X} \to b) = \Delta_\lambda(\mathcal{X} \to b) + \frac{1}{2}. \qquad (8)$$

We see that $\Delta_\lambda(\mathcal{X} \to b) = 0$ leads to $J(\lambda, \mathcal{X} \to b) = \frac{1}{2}$ which means no violation of the CHSH inequality. On the other hand, there is no assumption corresponding to freedom of choice regarding the outcomes, i.e., there are no physical grounds for assuming $J(\lambda \to B) = \frac{1}{2}$. Instead, one has

$$J(\lambda, \mathcal{X} \to B) = \Delta_\lambda(\mathcal{X} \to B) + J(\lambda \to B). \qquad (9)$$

Condition	Violation of CHSH possible?
$J(\lambda, \mathcal{X} \to b) = \frac{1}{2}$	No
$J(\lambda, \mathcal{X} \to B) = \frac{1}{2}$	No
$J(\lambda \to b) = \frac{1}{2}$	Yes ('freedom')
$J(\lambda \to B) = \frac{1}{2}$	Yes*
$\Delta_\lambda(\mathcal{X} \to b) = 0$	No
$\Delta_\lambda(\mathcal{X} \to B) = 0$	Yes*
SI: $P(A\|a, b, \lambda) = P(A\|a, \lambda)$	Yes**
OI: $P(A\|a, b, B, \lambda) = P(A\|a, b, \lambda)$	Yes**

TABLE I: The possibility of violation of the CHSH inequality in a local realistic model augmented with communication of $\mathcal{X}$ from Bob to Alice. $J(\lambda, \mathcal{X} \to b)$ and $J(\lambda, \mathcal{X} \to B)$ are the "guessed informations" by Alice, where λ denotes the hidden variables, and b and B are Bob's setting and outcome, respectively. $\Delta_\lambda(\mathcal{X} \to b)$ and $\Delta_\lambda(\mathcal{X} \to B)$ denote the "transmitted information" to Alice about Bob's setting and outcome, respectively, which is communicated via $\mathcal{X}$. (See main text for their definitions.) "No" in the right column means that the corresponding condition has to be violated to allow violation of the CHSH inequality. "Yes" means that there are models which satisfy the condition and violate the CHSH inequality. The starred "Yes*" in rows 4 and 6 indicate that for a violation either one of these conditions can hold, but not both. Similarly for the doubly-starred "Yes**" in rows 7 and 8, where for completeness we have included the previously known results in terms of SI and OI.

Thus, even if $\Delta_\lambda(\mathcal{X} \to B) = 0$, it is possible that $J(\lambda, \mathcal{X} \to B) > \frac{1}{2}$, if $J(\lambda \to B) > \frac{1}{2}$. Also $J(\lambda \to B) = \frac{1}{2}$ does not mean $J(\lambda, \mathcal{X} \to B) = \frac{1}{2}$ since $\Delta_\lambda(\mathcal{X} \to B)$ can be greater than 0. Summing up, neither $\Delta_\lambda(\mathcal{X} \to B) = 0$ nor $J(\lambda \to B) = \frac{1}{2}$ individually implies no violation of the CHSH inequality, although both of them together do. One can easily construct a toy model where $J(\lambda \to B) = \frac{1}{2}$ and violation occurs because the TI is $\Delta_\lambda(\mathcal{X} \to B) = \frac{1}{2}$ [17]. A toy model where $\Delta_\lambda(\mathcal{X} \to B) = 0$ and violation occurs is presented later. All different cases are presented in Table I.

To reinforce our conclusion that the freedom of choice assumption is responsible for the asymmetry, consider the possibility of 'superdeterminism', where everything is determined by the hidden variables, even the settings. In that case both $J(\lambda \to b) = 1$ and $J(\lambda \to B) = 1$, and consequently we have both $\Delta_\lambda(\mathcal{X} \to b) = 0$ and $\Delta_\lambda(\mathcal{X} \to B) = 0$: $\mathcal{X}$ is redundant as λ determines all there is to know. Settings and outcomes thus here appear on equal footing, and the conditions for violation of the CHSH inequality become identical for both. Indeed, only by giving up superdeterminism and allowing for freedom of choice for the settings we see the asymmetry between settings and outcomes arise. The assumption of freedom of choice of the settings enforces that, in order to get a violation of the Bell inequality, the message $\mathcal{X}$ *must* contain information about the setting, either implicit or explicit (though note that SI can be satisfied). It is however not needed that it carries information about the outcome.

Now we study explicit examples of non-local realistic models which violate the CHSH inequality. In all of them the non-

local information $\mathcal{X}$ is information about the distant setting.

Consider the model of Toner and Bacon [7]. One of the parties sends the bit $\mathcal{X} = \pm 1$ which is given by $\mathcal{X} = \mathrm{sgn}(\vec{b} \cdot \vec{\lambda}_1)\,\mathrm{sgn}(\vec{b} \cdot \vec{\lambda}_2)$, where $\vec{b}$ is a unit Bloch vector corresponding to Bob's setting and $\vec{\lambda}_1$ and $\vec{\lambda}_2$ are also unit vectors which play the role of hidden variables. The communication in this model can be compressed to $C \approx 0.85$ bits [7]. Exactly the same value is obtained for the mutual information between the bit sent and the setting, $I(\mathcal{X} : \vec{b}) = C$.

The model of [7] perfectly simulates all possible measurement results obtained on the singlet state. If one aims at simulation of the maximal violation of the CHSH inequality (with four fixed settings) allowed by quantum mechanics, then less communication is needed as shown in the following toy model with $\Delta_\lambda(\mathcal{X} \to B) = 0$. A binary random variable $\lambda = 0, 1$ is distributed from the source to Alice and Bob. His outcome for any choice of the setting is defined as $B = \lambda$, implying $J(\lambda \to B) = 1$ and thus $J(\lambda, \mathcal{X} \to B) = 1$ and $\Delta_\lambda(\mathcal{X} \to B) = 0$. If Bob's setting is $b = 0$ he sends always $\mathcal{X} = 0$, if his setting is $b = 1$ he sends $\mathcal{X} = 1$ with probability $p = \sqrt{2} - 1 \approx 0.414$, and $\mathcal{X} = 0$ otherwise. The outcome of Alice is given by $A = a\mathcal{X} \oplus \lambda$. In this model, the information content of $\mathcal{X}$ is 0.736 bits and again it is the information about the setting of Bob which is communicated. Note that classical players exchanging this amount of information achieve efficiency of quantum solutions to communication complexity problems and games based on the CHSH inequality [5].

In the Leggett-type [8] non-local model of Ref. [13], a real unit vector, i.e., an infinite number of bits, parameterizing the setting is being sent from one party to another, thus $\Delta_\lambda(\mathcal{X} \to b) > 0$. Note that this model violates SI, but obeys OI as it is deterministic [2].

Our last example is Bohm's theory [9]. Although here there is no explicit communication process, the information about the setting of the apparatus in one lab enters the formula for the velocity of the particle in the other lab. The analysis of the double Stern-Gerlach experiment shows that the velocity of one of the particles is given by [14] $v_1 = c_1 \tanh(c_2 \kappa)$, where κ is a parameter that describes the ratio between the magnetic field strengths at the two distant laboratories. The constants c_1 and c_2 do not depend on κ. Therefore, the local measurement outcome and the knowlege about the velocity *would* allow to infer the distant setting. Since tanh is an injective function, to determine v_1 all the bits defining κ have to be known to the mechanism that generates this velocity.

This last example shows that there need not be an actual communication process and our results are valid outside of the one-way communication paradigm. Indeed, it is irrelevant for our results how Alice obtained the information $\mathcal{X}$; one can think of it as extra information that tells about Bob's situation, and which is *somehow* available to Alice.

Conclusions— This work gives the general conditions which every non-local hidden-variable theory has to satisfy in order to allow for violation of the CHSH inequality. For there to be such a violation it must be the case that information about *both* the outcome and setting at one laboratory is *available* at the distant one, despite the fact that there is *no need* for both non-local setting and outcome dependence in the conditional (hidden) probabilities. The role of the setting is shown to be fundamentally different from that of the outcome and this asymmetry is shown to be due to the assumption of the experimenter's free setting choice. Because of this freedom the only way to learn a distant setting is to have non-local information transferral. By contrast, it is possible that the distant outcome can also be learnt from the shared hidden-variables, without any such non-local information transferral.

We use here the notion of availability of the non-local information in one laboratory about the setting chosen in the distant one. If under this term one would understand the existence of factual physical systems encoding this information, potentially allowing the information to be read, then the stronger results would follow: violation of Bell's inequality implies signalling at the hidden-variable level since the distant observer can (freely) choose his/her setting. This applies to both deterministic [15] as well as to stochastic models and goes clearly beyond that what can be concluded on the basis of an analysis using the conditions SI and OI.

We acknowledge support from the Austrian Science Foundation FWF within Project No. P19570-N16, SFB and CoQuS No. W1210-N16, the European Commission, Project QAP (No. 015848), the National Research Foundation and Ministry of Education in Singapore, and the Foundational Questions Institute (FQXi). The collaboration is a part of an ÖAD/MNiSW program.

[1] J. S. Bell, Physics **1**, 195 (1964).
[2] J. P. Jarrett, Noûs **18**, 569 (1984).
[3] A. Shimony, *Search for a Naturalistic World View*, vol. 2 (Cambridge University Press, Cambridge, UK, 1993).
[4] M. P. Seevinck, *Parts and Wholes*, arXiv:0811.1027.
[5] G. Brassard, Found. Phys. **33**, 1593 (2003); Č. Brukner *et al.*, Phys. Rev. Lett. **92**, 127901 (2004).
[6] B. Terhal *et al.*, Phys. Rev. Lett **90**, 157903 (2003).
[7] B. F. Toner and D. Bacon, Phys. Rev. Lett. **91**, 187904 (2003).
[8] A. J. Leggett, Found. Phys. **33**, 1469 (2003).
[9] D. Bohm, Phys. Rev. **85**, 166 (1952); P. R. Holland, *The Quantum Theory of Motion* (Cambridge University Press, Cambridge, UK, 1993).
[10] J.S. Bell, Epistemological Letters, vol. 9, March 1976.
[11] A. Fine, Phys. Rev. Lett. **48**, 291 (1982).
[12] J. F. Clauser *et al.*, Phys. Rev. Lett. **23**, 880 (1969).
[13] Supplementary information of S. Gröblacher *et al.*, Nature **446**, 871 (2007). Also at arXiv:0704.2529v2.
[14] D. Aerts *et al.*, Phys. Rev. A **73**, 034303 (2006).
[15] N.S. Jones *et al.*, Phys. Rev. A **71**, 042329 (2005).
[16] A. Valentini, Phys. Lett. A **297**, 273 (2002).
[17] A random binary hidden variable $\lambda = 0, 1$ is distributed to Alice and Bob. Bob's result for setting b is defined as $B = \lambda \oplus b$. Next, he communicates his outcome, $\mathcal{X} = B$. The result of Alice is given by $A = a(\mathcal{X} \oplus \lambda) \oplus \mathcal{X}$. Thus $J(\lambda, \mathcal{X} \to B) = 1$ and $J(\lambda \to B) = \frac{1}{2}$. Clearly, $A \oplus B = ab$, and the CHSH inequality is maximally violated. However, there is an *intrinsic setting infor-*

mation in this model as Alice can read the setting of Bob from the data available to her, $b = \mathcal{X} \oplus \lambda$; and thus $J(\lambda, \mathcal{X} \to b) = 1$ as well. This is what allows violation of the CHSH inequality.